

The Results of the NICE.II Iris Biometrics Competition

Kevin W. Bowyer

Department of Computer Science and Engineering

University of Notre Dame

Notre Dame, Indiana 46556 USA

kwb@cse.nd.edu

Abstract. The second phase of the Noisy Iris Challenge Evaluation attracted participation by 67 research groups from around the world. In contrast to all current commercial iris biometrics technology, the NICE competitions focus on performing iris biometrics on visible-light images. Whereas NICE.I focused on segmentation, NICE.II focused on performance in feature extraction and matching. The eight top-performing algorithms from NICE.II are considered, and suggestions are made for lessons that can be drawn from the results.

I. Introduction

Biometrics is typically defined as the study of methods of making measurements of physical, biological or behavioral attributes that can be used to identify a person. Within the field of biometrics, fingerprint, face and iris are often thought of as the current major general-purpose methods. This perception is reinforced by the fact that the Unique ID, or “Aadhaar”, project currently in progress in India aims to acquire face, ten-print fingerprint and both irises for all of the approximately 1.2 billion residents of India. At the time of this writing, the Unique ID project has already enrolled over 100 million persons [1]. For context, this is larger than the population of the United Kingdom, where plans for a biometric-enabled national identity card program were abandoned [2].

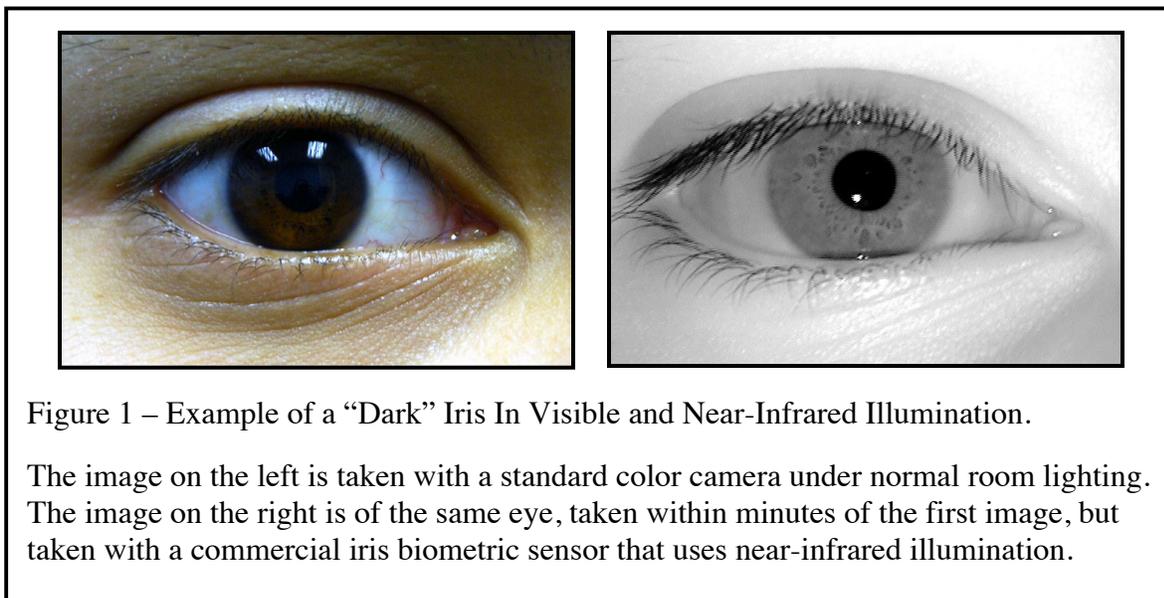
The NICE competition is concerned with methods that use the texture pattern of the iris as a means to recognize a person. Whereas commercial iris biometrics technology, such as that employed in India’s Unique ID program, uses near-infrared illumination of the eye, the NICE competition looks at what level of performance can be obtained using color images taken under already-existing illumination. The second phase of the NICE competition evaluated the performance of algorithms for feature extraction and matching.

II. A Brief Overview of the Current State of Iris Biometrics

Most researchers in the field of biometrics are aware that the development of the field has been hugely influenced by the work of John Daugman. While the basic concept patent for iris biometrics was awarded to Flom and Safir in 1987 [3], it was Daugman who introduced a compelling plan for practical implementation that was awarded a patent in 1994 [4]. Since that time, Daugman has introduced a number of improvements to his original approach (e.g., [5,6,7]).

Throughout the 1990s and into the early 2000s, the level of research activity in the field of iris biometrics was low. However, the volume of research activity in iris biometrics has increased dramatically in recent years. The first comprehensive survey of the iris biometrics research literature appeared in 2008 [8], listing approximately 180 references, and the number of published papers in the area has more than doubled since that time [9]. One of the trends that was identified in the 2008 survey [8], and which continues to be one major focus in iris biometrics research, is the desire for image acquisition under less controlled circumstances. There are various factors that enter into the concept of less controlled circumstances, and many of these are captured to some degree in the NICE competitions.

One element of less-controlled circumstances is the degree to which the user must actively cooperate with the sensor. The standard approach still predominates in commercial iris biometric technology – the user must stand still, look directly at the sensor, and actively cooperate with the sensor’s prompts to move closer to, or farther away from, the sensor. The “Iris On the Move” (IOM) system [10] developed by Sarnoff Corporation was the earliest to demonstrate that iris images of sufficient quality to be usable for iris biometrics could be obtained while the user was walking at normal speed and without explicitly posing. However, the near-infrared illuminators for this sensor were still located very close to the user, built into the frame of the portal that the user walks through. Also, if the user did not look forward at all during the moment of walking through the portal, there would be a failure to acquire. In general, if iris images are acquired while the person is “on the move” and without explicit cooperation in posing, the images may be less in focus, and have a more off-angle view of the iris, among possibly other complications.



The “active” illumination of the eye with near-infrared illuminators is another element in the concept of desiring less-controlled circumstances. The reason that near-infrared illumination is used for iris biometrics is that it allows the iris texture of both “light” and “dark” eyes to be well visualized. In effect, under appropriate near-infrared illumination, there are only “light” eyes with readily visualized iris texture. (See Figure 1.) If iris images are acquired under normal

visible light conditions, then the images may have highly variable contrast, are subject to less controlled specular highlights, and “dark” eyes may not have readily visualized iris texture, among other complications.

A third element in the concept of less-controlled circumstances is image quality. Iris biometric systems are typically filtering an input image stream to select a candidate image from which to compute an iris code. Candidate images are filtered based on image blur, degree of the iris that is occluded, and possibly other factors. When images are acquired under less controlled circumstances, it may not be feasible to filter out images that do not meet particular thresholds of image quality. There is more of an emphasis on making use of images that are of less than ideal quality.

The Iris Challenge Evaluation is a series of projects managed and conducted by the National Institute of Standards and Technology in the 2004-2007 time frame. Unlike the NICE competitions discussed in this paper, the ICE competitions used near-infrared illuminated iris images obtained from a standard commercial iris sensor. Self-reported results from ICE participants were presented at the ICE 2005 workshop [11] and results on sequestered data were presented in the ICE 2006 report [12]. As described on the NIST web pages for the ICE program [13], “The ICE 2005 was a technology development project for iris recognition. The ICE 2006 was the first large-scale, open, independent technology evaluation for iris recognition. The primary goals of the ICE projects were to promote the development and advancement of iris recognition technology and assess its state-of-the-art capability. The ICE projects were open to academia, industry and research institutes”. A dataset of nearly 70,000 iris images resulting from the ICE competitions is available to the research community [14].

III. The Noisy Iris Challenge Evaluation Competitions

The NICE competitions can be viewed as a reaction to the ICE program, specifically intended to focus on iris images acquired under visible light rather than under near-infrared illumination, and also to focus on “noise” of the type that occurs in more flexible image acquisition conditions. The NICE competitions got started with NICE.I in 2007 [15]. As explained in connection with NICE.I [16] - “It is highly probable that image capturing on less constrained conditions (either at-a-distance, on-the-move, with minor users' cooperation and within dynamic imaging environments) lead to the appearance of extremely heterogeneous images and with several other types of data in the captured iris regions (e.g., iris obstructions due to eyelids or eyelashes, reflections, off-angle or motion blurred images). For the terms of this contest, all these factors are considered as noise.”

The NICE.I competition focused on iris image segmentation. It required participants to provide an executable implementation that would process a visible-light eye image and produce a binary image that represents the iris region segmented from the rest of the image. Images used in the evaluation were manually given a classification of the noise-free / noisy iris pixels in each image. Based on comparison to the manual classification of image pixels, the pixel-level results of a segmentation algorithm were classified as true positive, true negative, false positive or false negative. A number of groups entered the NICE.I competition, and the best eight entries, in the sense of the lowest error rate for classifying pixels as noise-free iris or not, are described in a special issue [16].

The Noisy Iris Challenge Evaluation - Part II follows on the results of the NICE:I competition. As explained on the competition web page, “It also operates on iris images similar to the ones of the UBIRIS.v2” and “It is exclusively focused in the signatures encoding and matching stages of degraded visible wavelength iris images previously segmented, according to the segmentation method that outperformed in the segmentation contest (NICE:I)”. (The UBIRIS database of images is described in detail in [27].) Participants supplied an executable algorithm to the competition that takes as input two iris images and their corresponding segmentation masks. (See Figure 2.) The masks were created using the top-performing segmentation algorithm from the NICE:I competition. The result of the matcher for the competition is a dissimilarity metric. Algorithm performance was ranked based on the decidability value, d' , representing the difference in the means of the intra- and inter-class comparisons, divided by the square root of the average variance of the two classes. The NICE.II web pages list 67 algorithms registered for the competition. Papers describing the eight best-ranked algorithms are presented in this special issue. A brief overview of these algorithms is given in the next section.

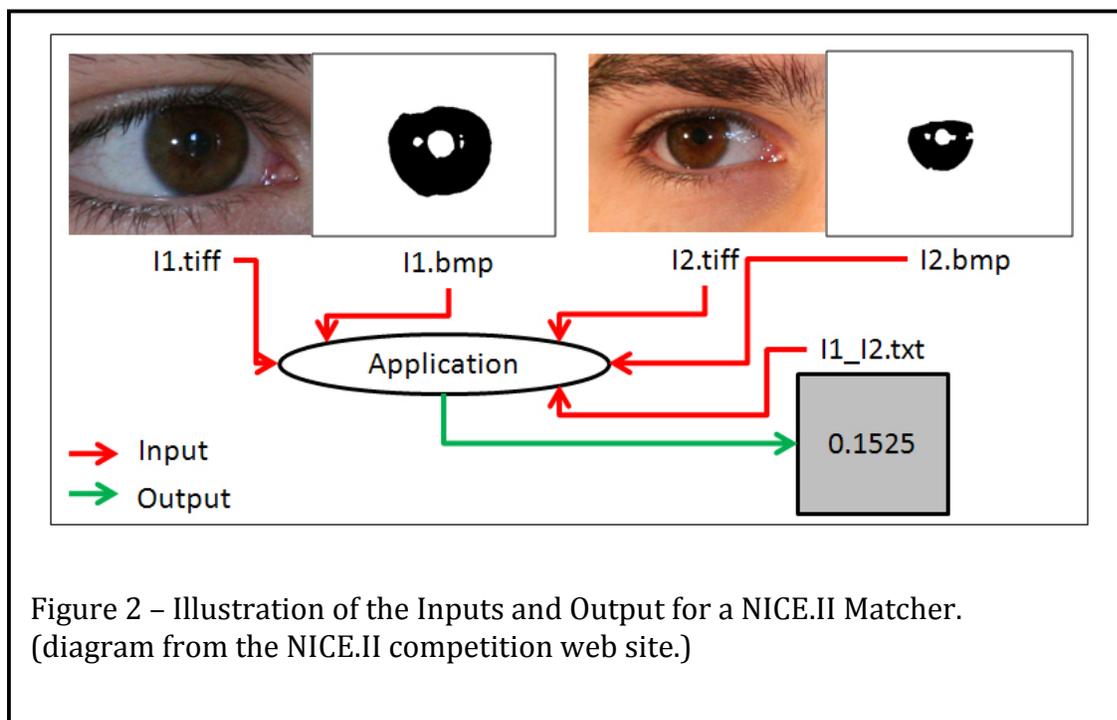


Figure 2 – Illustration of the Inputs and Output for a NICE.II Matcher. (diagram from the NICE.II competition web site.)

IV. The Best-Ranked Algorithms In the NICE.II Competition

The paper “Reliable Iris Recognition Algorithm Based on Reverse Biorthogonal Wavelet Transform” [18], by Robert Szewczyk, Kamil Grabowski, Malgorzata Napieralska, Wojciech Sankowski, Mariusz Zubert and Andrzej Napieralski from the Technical University of Lodz in Poland, describes the algorithm that earned eighth place in the NICE.II competition. Perhaps the most remarkable element of this approach is that the iris code used is only 384 bits in length. This is much smaller than the 2048-bit iris code used in the standard commercial implementation of Daugman’s approach. They first considered a range of possible wavelet decompositions for

use in analyzing the iris image. Based on looking at d' and equal error rate results from initial experiments, they selected the reverse biorthogonal 3.1 wavelet for use in their approach. Their segmentation algorithm generates a rectangular iris image of size 512x256. From the original color image, represented in red-green-blue color space, they create a monochrome image by averaging the red and green planes only. The 512x256 image is cropped to 256x256 “to remove influence of eyelashes on iris pattern and further decrease in computation of the signature encoding” and a histogram equalization step is performed. For the 256x256 image, the level-four wavelet decomposition results in 324 coefficients. A 324-bit binary code is created with a bit of the code set to 0 if the corresponding coefficient is less than the median of the coefficients, and 1 if it is greater. The difference between two 324-bit iris codes resulting from this approach is computed as the fractional Hamming distance between the codes.

The paper “Iris Recognition in Non-ideal Imaging Conditions” [19], by Peihua Li and Hongwei Ma from Heilongjiang University in China, describes the algorithm that earned seventh place. They estimate an ellipse for the limbic boundary of the iris region, and observe that this gives a more accurate boundary than a circular Hough transform. Their approach then creates a 64x256 rectangular image for the iris region. This image is considered as a set of non-overlapping sub-images of size 16x16. Matching of the iris codes for two images is then done as matching of the codes for the corresponding sub-images. Sub-images are matched allowing for a translation of the center of one sub-image within an 8x8 region relative to the other to minimize the Hamming distance. A search was performed over a set of candidate 2D Gabor filters to select the particular ones to be used in creating the iris code. The approach described in this paper is compared to Daugman’s 1993 algorithm – “We compare the proposed algorithm with the original Daugman’s algorithm (Daugman, 1993)” [19] – and found to perform better. However, Daugman’s 1993 algorithm would not be expected to be competitive with today’s state of the art; for example, it assumes circular limbic and pupillary boundaries. Segmentation using iris boundaries that are more general than circles has previously been described by a number of researchers [8].

The paper “Noisy Iris Recognition Integrated Scheme” [20], by Michele Nappi, Maria De Marsico and Daniel Riccio from the Università di Salerno in Italy, describes the algorithm that earned sixth place. Their approach to iris texture matching combines local binary patterns and discriminable textons, called “blobs”. They fit circles to the pupillary and limbic boundaries of the iris region and create a 40x360 rectangular iris image. The LBP and blob features are viewed as capturing complementary elements of the iris texture, with the LBP capturing “the textural regularities present in the iris, and blob identification for coding lighter or darker spots”. The iris region is considered as having five horizontal bands and a LBP histogram is computed for each band. This sequence of five LBP histograms then effectively becomes an “iris code” for the iris, though not obtained as, or structured as, a Daugman-like code. Dark and light blobs are found using a Laplacian-of-Gaussian operator at different scales and binarizing the result. Thus the iris code obtained for the blob analysis has a spatial structure that is somewhat more Daugman-like than the code for the LBP analysis. The name LBP-BLOB is given to the fusion of the two approaches, done by averaging the two scores. Results show that the combination of the two approaches performs better than either alone, and that the most useful scale for the blob analysis is different if the blob analysis is used alone versus used in combination with the LBP analysis.

The paper “Weighted Co-occurrence Phase Histogram for Iris Recognition” [21], by Peihua Li, Xiaomin Liu, and Nannan Zhao from Heilongjiang University in China, describes the

algorithm that earned fifth place. The co-occurrence histogram considers the phase angle at a given pixel along with that of each pixel at a distance of d pixels from the given pixel. Experiments are conducted with a training set to select the values of important parameters of the approach. Based on these experiments, the value of d is set at 2. Based on the parameter tuning experiments, it appears that the performance of the method decreases quickly with d greater than 3. So the co-occurrence relations captured are fairly local in nature. Also, with phase angle divided into 6 bins, the feature vector is of length $6^2 = 36$. Larger feature vectors did not result in any noticeable performance increase. The weighting aspect of the algorithm is that a pixel contributes to more than a single bin of the histogram, based on the difference between the phase value at the pixel and the range of near bins. The weighting factor is also tuned on the training set. The co-occurrence histogram is computed separately for each of a number of non-overlapping regions of the iris, with the intention to provide some robustness to alignment problems.

The paper “New Iris Recognition Method for Noisy Iris Images” [22], by Kwang Yong Shin, Gi Pyo Nam, Dae Sik Jeong, Dal Ho Cho, Byung Jun Kang, Kang Ryoung Park, and Jaihie Kim, from four different institutions in the Republic of Korea, describes the algorithm that earned fourth place. This approach employs an interesting sequence of classifier steps leading to a fusion of iris codes from the red, green and grayscale versions of the iris image. The first step uses cues from the eyelash distribution and specular reflections to classify the image as representing a left iris, a right iris, or undetermined. Eyelash density and specular reflection density both are generally greater near the medial canthus (corner of the eye near the nose) than near the lateral canthus (outer corner of the eye). Classification of the eye in the probe image is used to restrict the set of gallery images that are matched against. A second step uses differences between the two images’ iris region color distributions based on color spaces. The final step computes Hamming distances for iris codes based on each of the red and green color planes and the grayscale version of the image, and a weighted sum rule to fuse the distances. One interesting aspect of this approach is that its performance was good even though the left/right classification and color distribution classification steps sometimes made errors.

The paper “A Fusion Approach to Unconstrained Iris Recognition” [23], by Gil Santos and Edmundo Hoyle from the University of Beira Interior in Portugal, describes the algorithm that earned third place. Their approach is a multi-biometric fusion of five component algorithms. Both 1D wavelet and 2D wavelet zero-crossing maps are used as representations of iris texture. Another representation of iris texture used is based on the typical iris code resulting from 2D Gabor filters, but viewed as a binary “map” of the image and analyzed in a “comparison map” approach involving 16 sub-regions. For information extracted from the ocular region around the iris, two types of features are used. One is an LBP representation of the texture, and the other is based on a scale-invariant feature transform (SIFT) points. A logistic regression approach is used to fuse the five component results. Experiments were done to verify that the fusion result performs better than any of the individual components. Of the individual components, the zero-crossing based on 1D wavelets showed good performance, whereas the SIFT-based component showed the poorest performance.

The paper “Adaboost and Multi-Orientation 2D Gabor-based Accurate Noisy Iris Recognition” [24], by Qi Wang, Xiangde Zhang, Mingqi Li, Xiaopeng Dong, Qunhua Zhou and Yu Yin from Northeastern University in China, describes the algorithm that earned second place. They look at the concept of “noisy iris” partly in terms of whether or not the inner iris boundary

is accurately detected. The authors observe that, with their approach to iris segmentation, the outer boundary is generally localized well, but that the inner boundary can still be difficult to distinguish and localize. This work uses the standard rubber-sheet transform in cases where the inner and outer boundaries are both well detected, but uses a simpler rubber-sheet transform in cases where only the outer boundary is well detected. Part of the surprise in this work is that using a simplified rubber sheet transform based on only the outer iris boundary could lead to good performance. This work also considers iris codes generated for the whole iris and for local regions of the iris. The local regions are defined by considering top, right, bottom and left quadrants of the iris, and also either three or four concentric bands of the iris. Three bands are considered when both boundaries of the iris are used, and four bands are considered when only the outer boundary of the iris is used. Also, the Gabor filters are used at multiple orientations for the whole iris and for the local regions. Of the local regions considered, they find that features from the outer, lower region generally perform best. Features from regions of the iris nearest the pupil were not among the best-performing in either the case of a well-localized inner boundary or a not-well-localized inner boundary.

Finally, the paper “Noisy Iris Image Matching by Using Multiple Cues” [25], by Tieniu Tan, Xiaobo Zhang, Zhenan Sun, and Hui Zhang from the Chinese Academy of Sciences, describes the algorithm that placed first in the competition. Their approach is distinguished from most of the approaches described above by being a strongly multi-biometric approach that exploits multiple sources of information available in both the iris region and the surrounding ocular region. They use two iris-based sources of information: iris texture representation using ordinal measures as previously described by the CASIA research group, and iris color based matching. They also use two ocular-based sources of information: a texton-based matching of skin texture in the ocular region, and “semantic information” based on geometrical asymmetry of the eyelash distribution. Fusion of the four sources of matching information is done using a weighted sum of scores. The use of two sources of ocular information is an especially interesting aspect of this paper, as ocular biometrics has recently become a hot topic. Even more interesting is the use of two different types of ocular information, including one based on the eyelashes.

V. Discussion

The NICE.II competition has been quite successful overall, as evidenced in, for example, the large number of participants. It has focused attention on what can be achieved in iris biometrics using normal visible-light color images rather than the near-infrared illumination that is standard for current commercial systems. The d' value for the top-performing algorithm in NICE.II is 2.6. Relative to iris biometrics performed using near-infrared illumination under more controlled conditions, this is clearly modest performance. This result serves to highlight the challenging nature of the NICE.II problem, and suggests that there is still much room for possible improvement through future work.

Considering the papers in this special issue as a whole, several general themes are apparent for how to achieve improved performance in iris biometrics. One theme is fusion of multiple information sources. Using more than one type of texture feature extraction on the iris region and fusing the results improves performance relative to any one of the iris texture features. Using features extracted from the iris region and features extracted from the ocular region

improves results relative to using just the features from the iris region. The top-performing algorithm uses two sources of information from the iris region and two from the ocular region. Because the top-ranked algorithms seem to have relatively distinct technical approaches, it is likely that a fusion of the top algorithms would result in further performance improvement.

The fact that several of the top-performing algorithms use ocular features seems especially worth noting. The attention to ocular biometrics is relatively recent, with the topic having been popularized by Michael King's presentation at the 2007 *IEEE International Conference on Biometrics Theory, Applications and Systems*. The ocular features used in papers in this special issue go beyond simple texture features of the skin region around the eye. They include features related to the density of the eyelashes and to the density of specular highlights in the corners of the eye. These are feature sources that have not been widely exploited in ocular biometrics to date.

Another theme for improved performance is the fusion of results from multiple regions of the iris. This is sometimes motivated as a way to deal with inaccuracies in alignment, or a way to deal with inaccuracies in segmentation. In any case, computing feature vectors separately for parts of the iris region, matching them separately, and fusing the results has been used by several research groups as a means to improve performance.

Another theme that is evident applies specifically to the use of color images. The authors that explicitly considered the use of red, green and blue components of a color image seem to uniformly agree that the blue component of a red-green-blue color image is not useful for iris biometrics.

Several of the papers in this special issue mention the "lifetime stability" of the iris. It was once believed that iris texture was stable throughout adult life in a way that meant that there was no template aging effect. Mansfield and Wayman [26] define template aging as – "Template ageing refers to the increase in error rates caused by time related changes in the biometric pattern, its presentation, and the sensor". Recent studies have shown that iris biometrics is in fact subject to a template aging effect. For example, Fenker and Bowyer [27] worked with a dataset involving two years of images for 86 irises and found that "there is an increase in false reject rate across the range of potential decision threshold values."

All researchers interested in iris biometrics should find something useful and valuable in the results of the NICE competitions and in this collection of papers detailing the top-performing algorithms in NICE.II.

References.

- [1] “Aadhaar gains momentum, clocks two crores in October”, *The Economic Times*, November 2, 2011.
- [2] “Identity cards scheme will be axed 'within 100 days'”, BBC News, 27 May 2010, http://news.bbc.co.uk/2/hi/uk_news/politics/8707355.stm
- [3] L. Flom and A. Safir, “Iris Recognition System”, US Patent # 4,641,349, February 3, 1987.
- [4] J. G. Daugman, “Biometric Personal Identification System Based on Iris Analysis”, US Patent # 5,291,560, March 1, 1994.
- [5] J. Daugman, “How Iris Recognition Works”, *IEEE Transactions on Circuits and Systems for Video Technology* 14 (1), Jan. 2004, 21-30.
- [6] J. Daugman, “Probing the Uniqueness and Randomness of IrisCodes: Results From 200 Billion Iris Pair Comparisons”, *Proceedings of the IEEE* 94 (11), Nov. 2006, 1927-1935.
- [7] J. Daugman, New Methods in Iris Recognition *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 37 (5), Oct. 2007, 1167-1175.
- [8] K. W. Bowyer, K. Hollingsworth, and P. J. Flynn, “Image Understanding for Iris Biometrics: A Survey”, *Computer Vision and Image Understanding* 110 (2008) 281–307.
- [9] K. W. Bowyer, K. P. Hollingsworth and P. J. Flynn, A Survey of Iris Biometrics Research: 2008-2010, to appear.
- [10] J. R. Matey, O. Naroditsky, K. Hanna, R. Kolczynski, D. J. LoIacono, S. Mangru, M. Tinker, T. M. Zappia, and W. Y. Zhao, “Iris on theMove: Acquisition of Images for Iris Recognition in Less Constrained Environments”, *Proceedings of the IEEE* 94 (11), November 2006, 1936-1947.
- [11] P.J. Phillips, K.W. Bowyer, P.J. Flynn, X. Liu and W.T. Scruggs, “The Iris Challenge Evaluation 2005”, *IEEE International Conference on Biometrics: Theory, Applications and Systems*, September 2008. Digital Object Identifier: 10.1109/BTAS.2008.4699333
- [12] P. J. Phillips, W. T. Scruggs, A. O'Toole, P. J. Flynn, K. W. Bowyer, C. L. Schott and M. Sharpe, “FRVT 2006 and ICE 2006 Large-Scale Experimental Results”, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32 (5), May 2010, 831-846.
- [13] National Institute of Standards and Technology (NIST), ICE – Iris Challenge Evaluation, <http://www.nist.gov/itl/iad/ig/ice.cfm>
- [14] Kevin W. Bowyer and Patrick J. Flynn, The ND-IRIS-0405 Iris Image Dataset, www.nd.edu/~cvrl/CVRL/Data_Sets.html
- [15] H. Proenca and L.A. Alexandre, “The NICE.I: Noisy Iris Challenge Evaluation - Part I”, *IEEE International Conference on Biometrics: Theory, Applications, and Systems*, September 2007. Digital Object Identifier: 10.1109/BTAS.2007.4401910

- [16] “Special Issue on Segmentation of Visible Wavelength Iris Images Captured At-a-distance and On-the-move”, *Image and Vision Computing* 28 (2), February 2010.
- [17] H. Proença, S. Filipe, R. Santos, J. Oliveira, L. A. Alexandre, “The UBIRIS.v2: A Database of Visible Wavelength Iris Images Captured On-The-Move and At-A-Distance”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32 (8), pp. 1529-1535, August, 2010.
- [18] R. Szewczyk, K. Grabowski, M. Napieralska, W. Sankowski, M. Zubert and A. Napieralski, “Reliable Iris Recognition Algorithm Based on Reverse Biorthogonal Wavelet Transform” *Pattern Recognition Letters*, to appear.
- [19] P. Li and H. Ma, “Iris Recognition in Non-ideal Imaging Conditions” *Pattern Recognition Letters*, to appear.
- [20] M. Nappi, M. De Marsico and D. Riccio, “Noisy Iris Recognition Integrated Scheme”, *Pattern Recognition Letters*, to appear.
- [21] P. Li, X. Liu, N. Zhao, “Weighted Co-occurrence Phase Histogram for Iris Recognition”, *Pattern Recognition Letters*, to appear.
- [22] K. Y. Shin, G. P. Nam, D. S. Jeong, D. H. Cho, B. J. Kang, K. R. Park, J. Kim, “New Iris Recognition Method for Noisy Iris Images”, *Pattern Recognition Letters*, to appear.
- [23] G. Santos and E. Hoyle, “A Fusion Approach to Unconstrained Iris Recognition”, *Pattern Recognition Letters*, to appear.
- [24] Q. Wang, X. Zhang, M. Li, X. Dong, Q. Zhou, Y. Yin, Adaboost and multi-orientation 2D Gabor-based accurate noisy iris recognition, *Pattern Recognition Letters*, to appear.
- [25] T. Tan, X. Zhang, Z. Sun, H. Zhang, “Noisy Iris Image Matching by Using Multiple Cues”, *Pattern Recognition Letters*, to appear.
- [26] A.J. Mansfield and J.L. Wayman, Best Practices in Testing and Reporting Performance of Biometric Devices, Version 2.01, Aug 2002.
- [27] S. Fenker and K. W. Bowyer, “Experimental evidence of a template aging effect in iris biometrics,” *IEEE Computer Society Workshop on Applications of Computer Vision*, January 2011.